Unfolded conformations of α-lytic protease are more stable than its native state

Julie L. Sohl*†‡, Sheila S. Jaswal‡§ & David A. Agard*§

** Graduate Group in Biophysics and the Howard Hughes Medical Institute, § Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143-0448, USA*
‡ These authors contributed equally to this work
† Present address: Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, California 94720, USA.

α-Lytic protease (αLP), an extracellular bacterial protease, is synthesized with a large amino-terminal pro-region that is essential for its folding *in vivo* and *in vitro*[1,2]. In the absence of the pro-region, the protease folds to an inactive, partially folded state, designated 'I'. The pro-region catalyses protease folding by directly stabilizing the folding transition state (>26 kcal mol^{-1}) which separates the native state 'N' from I[1,3]. Although a basic tenet of protein folding is that the native state of a protein is at the minimum free energy[4], we show here that both the I and fully unfolded states of αLP are lower in free energy than the native state. Native αLP is thus metastable: its apparent stability derives from a large barrier to unfolding. Consequently, the evolution of αLP has been distinct from most other proteins: it has not been constrained by the free-energy difference between the native and unfolded states, but instead by the size of its unfolding barrier.

The αLP N state is compact, with a well ordered hydrophobic core, and is stable to chemical and thermal denaturation (see ref. 5 and below). In contrast, the I state has some secondary structure but little or no tertiary structure[1]. Gel filtration[1] and analytical ultracentrifugation (Fig. 1a) indicate that I is a greatly expanded monomer. Not surprisingly, I is temperature-sensitive[1] and it is stabilized by less than 1 kcal mol^{-1} relative to the unfolded state 'U' (Fig. 1b). The conventional view of the forces that drive protein folding[4] would place N at the minimum Gibbs free energy. In principle, however, it is possible for a protein to function even if its native state is not at the free-energy minimum, provided that the kinetic barrier is sufficient to prevent unfolding over the protein's functional lifetime. We now demonstrate that this is the case for αLP.

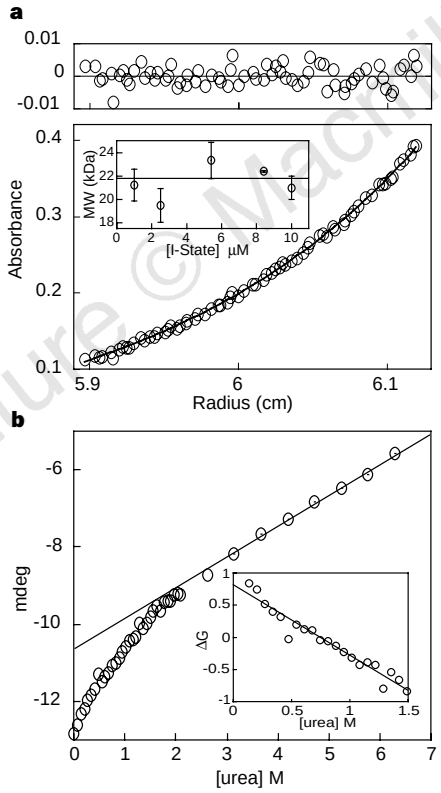

Figure 1 Characterization of the αLP I state. **a**, Sedimentation equilibrium analysis of 10 μM αLP I state with monomer fit and residuals. As shown in the inset, the I state remains monomeric, with an average M_r of 21.8K over at least a 10-fold concentration range. **b**, Urea denaturation of the I state followed by its circular dichroism signal at 225 nm. Inset shows the free energies for the I ↔ U transition as a function of urea, calculated using the displayed linear unfolded baseline and by assigning the I-state baseline to the 0 M urea value. The calculated $\Delta G_{I \leftrightarrow U}$ at 0 M urea must be considered a maximum value.

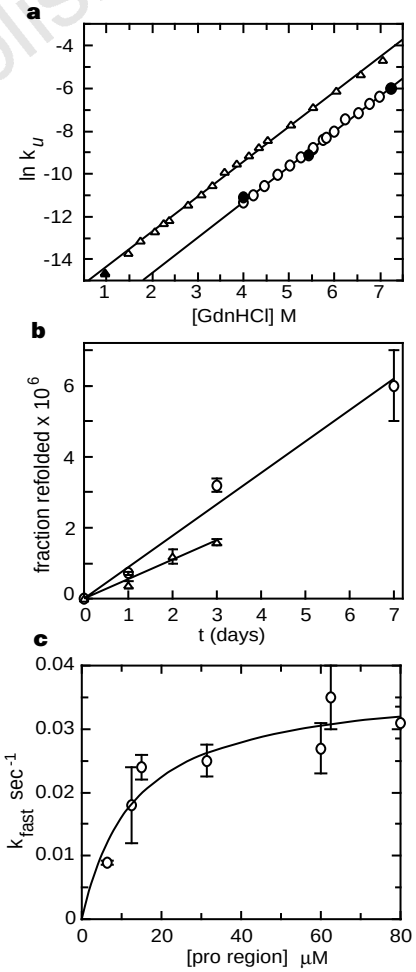

Figure 2 The rates of αLP unfolding and refolding. **a**, The logarithm of the unfolding rates at 4 °C (circles) and 25 °C (open triangles) as a function of denaturant (guanidinium (Gdn) hydrochloride) concentration. Open symbols denote the rates of tertiary structure loss of S195A observed using tryptophan fluorescence; the filled triangle represents the rate of activity loss of the wild-type protease; filled circles show the rates of secondary structure loss in S195A monitored by circular dichroism at 225 nm. **b**, Pro-region-independent folding of the I to N states as a function of time at 4 °C (circles) and 25 °C (triangles). **c**, The concentration dependence of the fast phase rates of pro-region-catalysed I to N folding.

The large folding barrier prevents equilibration of I and N on any practical time scale; thus the relative stability of these states is addressed indirectly through the rates of folding (k_f) and unfolding (k_u). At 25 °C, unfolding of the protease in 1 M guanidinium-HCl can be monitored by loss of activity. With increasing denaturant concentration, measurements of wild-type protease unfolding are complicated by autolysis, so we used a variant of the protease in which the active-site serine is replaced by alanine (Ser 195 → Ala, S195A, where the residue numbering for αLP is based on homology to chymotrypsin[3]) to measure the rates of unfolding as a function of denaturant at 4 °C and 25 °C. During unfolding, secondary and tertiary structure are lost simultaneously and the spectroscopic data are best fitted by a single exponential function of time. The observed rates are plotted as a function of guanidinium-HCl concentration in Fig. 2a. By linear extrapolation to zero concentration of guanidinium, the rate of unfolding is $1.8 \times 10^{-8} \pm 0.2 \times 10^{-8}\,s^{-1}$ at 4 °C ($t_{1/2} = 1.2$ y) and $1.2 \times 10^{-7} \pm 0.6 \times 10^{-7}\,s^{-1}$ at 25 °C ($t_{1/2} = 70$ d). Urea denaturation at 4 °C yields the same extrapolated rate within error (data not shown), supporting the use of a linear extrapolation to determine k_u.

To measure k_f, we incubated solutions of I at 4 °C and 25 °C and assayed aliquots for N-state protease activity as a function of time. The very small fraction of I that refolds during the incubation was detected with an assay of higher sensitivity than that used previously[1] (Fig. 2b). Based on the concentration of refolded N and the total protease concentration, the initial rate of folding k_f was found to be $1.18 \times 10^{-11} \pm 0.06 \times 10^{-11}\,s^{-1}$ at 4 °C ($t_{1/2} > 1{,}800$ y) and $6.0 \times 10^{-12} \pm 0.4 \times 10^{-12}\,s^{-1}$ at 25 °C ($t_{1/2} > 3{,}600$ y). These rates are significantly slower than the rates of unfolding. The equilibrium free energy, calculated from the ratio of k_f to k_u, favours I by 4.0 ± 0.1 kcal mol^{-1} at 4 °C and by 5.9 ± 0.5 kcal mol^{-1} at 25 °C. Therefore, the I state of αLP, not the N state, is at the minimum free energy over a broad range of temperatures.

αLP folds efficiently to its metastable native conformation only with the assistance of its pro-region. The catalysed folding reaction has two time constants, with the population of the fast phase being roughly twice that of the slow phase. The slow phase, which is essentially independent of pro-region concentration, has a time constant of ~200 s (data not shown), and may be a consequence of proline isomerization, as αLP has one *cis* and three *trans* prolines. The fast phase rates are plotted as a function of pro-region concentration in Fig. 2c. From these data, the pro-region is calculated to bind I with 13 ± 5 μM affinity and the folding rate is calculated to be $0.037 \pm 0.004\,s^{-1}$.

With the data shown in Figs 1 and 2, a free-energy diagram for the protease-folding reaction in the presence and absence of pro-region can be constructed (Fig. 3). Remarkably, both I and U are more stable than the folded N state. The pro-region shifts the energetics thermodynamically to favour the N·pro-region complex over the I·pro-region complex, thereby providing the driving force that brings the protease into its native, active conformation. Proteolytic degradation of the pro-region then 'locks' αLP in its metastable active N state. The pro-region is much more than a native-state template. It actively catalyses the folding reaction by stabilizing the folding transition state. The pro-region binds the folding transition state more tightly than I or N, and mutants of pro-region have been identified that preferentially alter this transition-state stabilization[5].

Investigations into the enthalpic and entropic contributions to the free energy difference (ΔG) between I and N provide insight into the unusual stability of the unfolded states of αLP. Because direct measurement of the enthalpy difference (ΔH) between I and N is not possible, we measured ΔH by titration calorimetry using the thermodynamic cycle shown in Fig. 4. At 10 °C, the native conformation is favoured enthalpically by 18 ± 1.5 kcal mol^{-1} over I. Therefore, the stability of I must be entropic in origin.

Comparison of αLP with the homologous but thermodynamically stable protease, chymotrypsin, reveals an intriguing potential source of excess entropy. The αLP sequence contains 16% glycine compared to only 9% in chymotrypsin. Glycines lack a side chain, which increases the number of conformations accessible to unfolded states. The ten additional glycines in αLP are predicted to contribute an additional ~7 kcal mol^{-1} of configurational entropy to the unfolded state compared to that of chymotrypsin at 4 °C (ref. 6). Removal of this entropic source alone would be sufficient to place the N state at the global free-energy minimum for αLP.

An unusually low conformational entropy of the αLP N state may also contribute to increasing the unfolding entropy. Although native states are often dynamic, the N state of αLP is quite rigid: the protease is not readily digested by itself or by other proteases, it has ~40 core amides with >10^{10} protection factors, and the B factors are unusually low (J. Davis, J.L.S. and D.A.A., unpublished results, and ref. 3). Some of the rigidity may arise from the fact that the

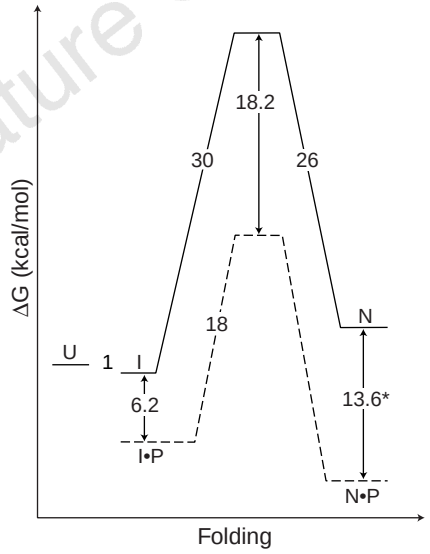

Figure 3 Free energy of αLP folding with and without its pro-region at 4 °C. Using the pro-region I-state binding constant (Fig. 2c), the pro-region N-state inhibition constant[17] and the rates of folding and unfolding (Fig. 2a, b) in combination with transition state theory, the free-energy profile of αLP folding is drawn. Dashed lines denote energetic changes resulting from the addition of pro-region. The free-energy difference between I and U is not precisely known but is estimated to be less than ~1 kcal mol^{-1} (see Fig. 1b). Asterisk indicates that the pro-region binding of the native state was measured at 25 °C and pH 8.

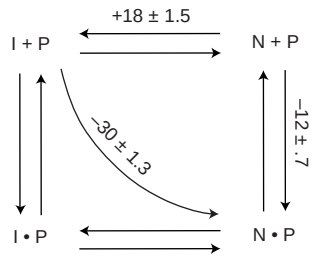

Figure 4 Dissection of the free energy difference between I and N. The enthalpy difference between I and N was determined at 10 °C using the outlined thermodynamic cycle. The enthalpies of I + P to N·P and N + P to N·P were measured. The total enthalpy change around the cycle must equal zero. Therefore, the enthalpy of the N- to I-state transition is +18 ± 1.5 kcal mol^{-1} at 10 °C.

loops in αLP are generally shorter and are likely to be less flexible than those found in chymotrypsin. Many of the glycines unique to αLP are also found in these loops and so may reduce the entropy of the N state while increasing the configurational entropy of the I state.

The rigidity of the native state of αLP may be the product of evolutionary pressures to suppress autolysis and thereby extend the lifetime of the protease, which provides nutrients for its host. The unfavourable and presumably highly cooperative unfolding barrier of αLP would be expected to reduce sources of proteolytic sensitivity such as local breathing motions and/or partial unfolding of the native state. Similarly large kinetic barriers are likely to be present in most extracellular bacterial proteases; virtually all are synthesized with pro-regions and, where examined, the pro-regions catalyse protease folding[7–9]. There is a strong correlation between high glycine content and the presence of a pro-region. As the average glycine content of αLP and other chymotrypsin-like proteases made with pro-regions is 18%, compared with 9% in family members that do not contain a pro-region, incorporating more glycines may be one mechanism by which large kinetic barriers and pro-region catalysts have co-evolved in extracellular bacterial proteases in order to prolong their functional lifetime.

Large barriers have been observed in the folding of several other proteins that are not synthesized with pro-regions. The kinetic barriers of haemaglutinin, luciferase, the serpin PAI-1, and of the prion protein PrP, all function to enhance the stability of one compact native-like state relative to another[10–13]. In contrast, αLP has evolved a kinetic barrier to isolate its native state from unfolded states. This has enabled the protease to optimize its functional properties independently of the thermodynamic consequences, resulting in a kinetically stable but thermodynamically unstable native state. □

Methods

Protein expression and purification. αLP and pro-region were prepared according to published protocols[14]. The construct for the inactive mutant (S195A) has been described[14]. Expression and purification of S195A was as described[15], except that the medium contained 1.5% yeast extract, 1% NaCl, 60 mM ACES, pH 6.3, and induction was at 12 °C. Pro-region was digested from the non-covalent S195A–pro-region complex with trypsin-coupled beads at pH 7 or with pepsin at pH 3. The S195A X-ray structure is virtually identical to that of wild-type protease (S.S.J., S. Rader and D.A.A., unpublished results).

Analytical ultracentrifugation. Sedimentation equilibrium experiments at 4 °C were carried out on I state in 10 mM potassium acetate, pH 5, on a Beckman Model XLA analytical ultracentrifuge at speeds of 18,000 and 22,000 r.p.m., scanning at 230 nm and 280 nm. Data were evaluated using XL-A data analysis software and fitted best at all concentrations to the single non-associating species model. Global analysis of nine data sets gave a relative molecular mass (M_r) of 21.7K with a 95% c.i. of ±0.9K.

Urea denaturation. The stability of I state was measured as a function of urea at 225 nm in an AVIV 62DS circular dichroism spectropolarimeter at 4 °C. 3.5 μM I state in 10 mM potassium acetate at pH 5 was titrated with 4 M and 8 M urea. The unfolded baseline was calculated by a linear fit of the 3–6 M urea data.

Wild-type protease unfolding. 4 μM protease in 0.98 M guanidinium-HCl, 10 mM potassium acetate, pH 5, was incubated at 25 °C. Activity was measured 13 times on duplicate aliquots according to published protocols over 80 days[1].

S195A unfolding. Reactions contained S195A in 10 mM potassium acetate, pH 5, at concentrations of 0.1–1.75 μM (fluorescence) or 7 μM (circular dichroism). Fluorescence measurements were made in a 8100SLM-Aminco fluorimeter. Circular dichroism spectra were monitored on a Jasco J-710 spectropolarimeter. The error in the 25 °C rate constant was determined from the difference between the linear extrapolation and denaturant binding model[16] fits.

Refolding from the I state to the N state. 2.4–6.5 μM I state in 10 mM potassium acetate, pH 5, was incubated at 4 °C and 25 °C. Initially and at each timepoint, 3-ml aliquots of I state were treated with 50 μg pepsin at pH 2.5, which digests the I state but leaves native protease intact, then concentrated. The protease was assayed by following the absorbance change at 324 nm in a solution containing 1 mM succinyl-Ala-Ala-Pro-Ala-thiobenzyl ester (Enzymes Systems Products), 250 μM aldrithiol-4 (Sigma), 2.5% DMSO and 0.1M Tris, pH8. Addition of 1 μM pro-region inhibited the observed activity, indicating that activity is due to αLP N state. The concentration of native protease, $[N]$, was determined from a standard curve. The detection limit is ~6.0 fmolar. Intact I state, $[I]_0$, was measured at each timepoint, as it decreases as a function of time owing to *I*-state proteolysis by refolded N.

Data from three experiments were fitted simultaneously to determine k_f using data in which the amount of intact I state was within 10% of the starting value. Calculations are based on a reversible equilibrium, $I \leftrightarrow N$. The rate equation for this equilibrium is $[I]_t/[I]_0 = [1/(k_f + k_u)] \times [k_f\{\exp - (k_f + k_u)t\} + k_u]$, where the concentration of I state at time t is given by $[I]_t = [I]_0 - [N]$, assuming $[I]_t = [I]_0$ at $t = 0$. As both $k_f t$ and $k_u t$ are small, the Taylor series approximation is applied, yielding: $k_f = ([N]/[I]_0)/t$.

Refolding in the presence of the pro-region. Pro-region-catalysed refolding of the I state was measured at 4 °C in 20 mM potassium phosphate at pH 7.2 (ref. 1). The pro-region:I state concentration ratio was maintained at ≥15:1 and reaction rates were fitted by a five-parameter double exponential. The plot of the fast phase rates as a function of pro-region concentration was analysed by assuming a fast I + P ↔ 1·P pre-equilibrium followed by a slow, concentration-independent I·P → N·P transition.

Titration calorimetry. Enthalpies were measured in 20 mM potassium acetate, 42 mM guanidinium-HCl, pH 5, buffer using a Microcal Omega titration calorimeter. Pro-region (~100 μM) was titrated with I or N from 8–12 °C. The enthalpies at 10 °C are the average of nine measurements. Temperatures below 8 °C and above 12 °C were not used owing to low signal-to-noise and complications from pro-region unfolding, respectively.

Received 15 May; accepted 30 July 1998.

1. Baker, D., Sohl, J. L. & Agard, D. A. A protein-folding reaction under kinetic control. *Nature* **356,** 263–265 (1992).
2. Silen, J. L. & Agard, D. A. The α-lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*. *Nature* **341,** 462–464 (1989).
3. Fujinaga, M., Delbaere, L. T., Brayer, G. D. & James, M. N. Refined structure of α-lytic protease at 1.7 Å resolution. Analysis of hydrogen bonding and solvent structure. *J. Mol. Biol.* **184,** 479–502 (1985).
4. Makhatadze, G. I. & Privalov, P. L. Energetics of protein structure. *Adv. Protein Chem.* **47,** 307–425 (1995).
5. Peters, R. J. *et al.* Pro region C-terminus: Protease active site interactions are critical in catalyzing the folding of α-lytic protease. *Biochemistry* **37,** 12058–12067 (1998).
6. D'Aquino, J. A. *et al.* The magnitude of the backbone conformational entropy change in protein folding. *Proteins* **25,** 143–156 (1996).
7. Baker, D., Shiau, A. K. & Agard, D. A. The role of pro regions in protein folding. *Curr. Opin. in Cell Biol.* **5,** 966–970 (1993).
8. Eder, J. & Fersht, A. R. Pro-sequence-assisted protein folding. *Mol. Microbiol.* **16,** 609–614 (1995).
9. Bryan, P. *et al.* Catalysis of a protein folding reaction: mechanistic implications of the 2.0 Å structure of the subtilisin–prodomain complex. *Biochemistry* **34,** 10310–10318 (1995).
10. Clark, A. C. *et al.* Kinetic mechanism of luciferase subunit folding and assembly. *Biochemistry* **36,** 1891–1899 (1997).
11. Chen, J. *et al.* A soluble domain of the membrane-anchoring chain of influenza virus hemagglutinin (HA2) folds in *Escherichia coli* into the low-pH-induced conformation. *Proc. Natl Acad. Sci. USA* **92,** 12205–12209 (1995).
12. Harrison, P. M., Bamborough, P., Daggett, V., Prusiner, S. B. & Cohen, F. E. The prion folding problem. *Curr. Opin. Struct. Biol.* **7,** 53–59 (1997).
13. Wang, Z., Mottonen, J. & Goldsmith, E. J. Kinetically controlled folding of the serpin plasminogen activator inhibitor 1. *Biochemistry* **35,** 16443–16448 (1996).
14. Sohl, J. L., Shiau, A. K., Rader, S. D., Wilk, B. J. & Agard, D. A. Inhibition of α-lytic protease by pro region C-terminal steric occlusion of the active site. *Biochemistry* **36,** 3894–3902 (1997).
15. Mace, J. E. & Agard, D. A. Kinetic and structural characterization of mutations of glycine 216 in α-lytic protease: a new target for engineering substrate specificity. *J. Mol. Biol.* **254,** 720–736 (1995).
16. Myers, J. K., Pace, C. N. & Scholtz, J. M. Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4,** 2138–2148 (1995).
17. Baker, D., Silen, J. L. & Agard, D. A. Protease pro region required for folding is a potent inhibitor of the mature enzyme. *Proteins* **12,** 339–344 (1992).

Acknowledgements. We thank P. Walter and P. Peluso for use of their fluorimeter, S. Marqusee for use of her Aviv CD and titrator, Y.-R. Yang for assistance with analysis of data from analytical ultracentrification, N. Sauter for assistance with figures, and A. Derman, M. Gustafsson, T. Mau, A. Shiau and H. S. Chan for comments on the manuscript. This work was supported by the Howard Hughes Medical Institute. S.S.J. is supported by a Howard Hughes Medical Institute predoctoral fellowship. J.L.S. was supported by NIH training grants.

Correspondence and requests for materials should be addressed to D.A.A. (e-mail: agard@msg.ucsf.edu).